Lecture 3

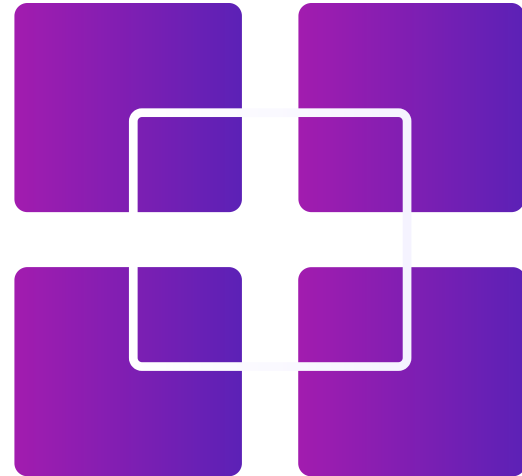**- Field Experiments**
- Motivation

Agenda

- Method: Field experiments
    - Two NYT articles
    - Working with organizations
    - Practice: How to design a field experiment
- Topic: Motivation
    - History of motivation research
    - Field experiments in motivation: Milkman et al.
    - Discussion questions
- Next class

Method: Field experiments

Two NYT articles: Field experiments gone bad vs. well.

- Restaurant study

- Professor study

Two NYT articles: Field experiments gone bad vs. well.

- What was the problem?

Two NYT articles: Field experiments gone bad vs. well.

- What was the problem?
- What is the difference between the two cases?
  - Consent; deception; debriefing
  - Consequences in real world
  - Implications of research (e.g., COVID vaccine vs. lip Botox)
- Who makes the final call? IRB (next week)

Two NYT articles: Field experiments gone bad vs. well.

- Restaurant study
- Let's try to design one by ourselves

Working with organizations (Step by step) 1

- Start the project
  - Take the organization's perspective
  - Communication is the key!
  - Get things in writing (signed contract ideally, in advance)

Working with organizations (Step by step) 2

- Plan an experiment
    - Which data to collect (be mindful of confidentially)
    - What data organizations are reluctant to share?
    - What data organizations are willing to share?
    - Be aware of their timeline (e.g., labor union negotiation)

Working with organizations (Step by step) 3

- Design an experiment
  - Want to be as close to lab as possible
  - Random assignment in field (e.g., morning, summer)
  - Who will participate? (e.g., which division)
  - Is it natural behavior? (e.g., hand washing)
  - How will you measure?

Working with organizations (Step by step) 4

- Conduct an experiment
  - Have control over the process
  - Who will be there? (e.g., experimenter vs. boss)
  - Debriefing both participants and organization

Practice: How to design a field experiment

○ Q. "Does having a family photo on your desk improve your motivation?"

Topic: Motivation

History for motivation research

- Goal setting theory (Locke & Latham, 2006)
- Goal systems theory (Kruglanski, 2003)
- Expectancy theory (e.g., Effort – Performance – Reward)
- Equity theory (e.g., input vs. output)

Field experiments in motivation

Milkman et al.

Holding the Hunger Games Hostage at the Gym

An Evaluation of Temptation Bundling

Milkman, Minson, and Volpp (2014)

Motivation: bundle "Want" with "Should" behaviors

How can we engage in healthy behaviors?

Professor Katherine Milkman loves to read about and listen to "addictive" fiction novels like The Hunger Games. So a few years ago, when she was having trouble making it to the gym on a regular basis, she decided to enjoy these "wanted" audiobooks only when she was exercising.

Surprisingly, Milkman began hitting the gym five days a week. The fact makes her consider about such self-commitment strategy:

bundling "want" experience with "should" behaviors to promote "should" behaviors
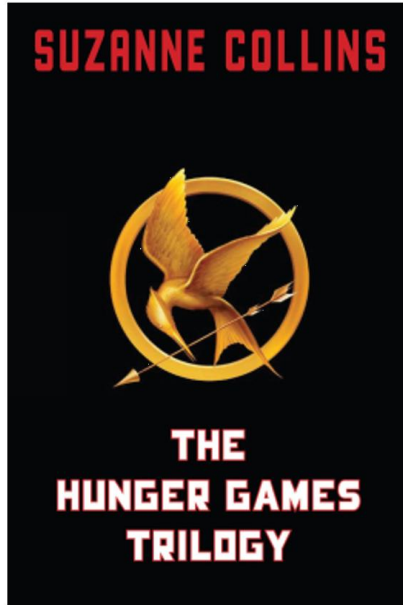
"Want" Experience × "Should" Experience

Temptation Bundling

Research Focus

Topic

introduce and evaluate the effectiveness of "temptation bundling" method which improves self-control problems

Method

A field experiment to examine whether:

- such bundles can increase "should" behaviors
- people would pay for these restrictive bundles

"Want"	Self-control Problems	"Should"
Instantly gratifying activities e.g. watching the TikTok videos, online surfing in Twitter	People indulge in "want" behaviors but ignore the "should" behaviors	**Beneficial behaviors but with delayed rewards** e.g. exercising, not staying up late

Previous Research

Self-control problems and behavioral interventions

○ incentivize *repeated* gym attendance can produce long-lasting effect (Charness and Gneezy 2009, Acland and Levy 2013)

○ **Willpower**: failure in self-control suggests limited willpower (see, e.g., Della Vigna and Malmendier 2006)

○ **Less should behaviors**: people intend to exercise and diet tomorrow but frequently lack the necessary willpower to act on those good intentions today (Della Vigna and Malmendier 2006, Milkman et al. 2009, Read and van Leeuwen 1998, Royer et al. 2012)

○ **Regret**: limited willpower makes it difficult for individuals to resist temptation, which induces regret after the fact (for a review, see Milkman et al. 2008).

How does Temptation Bundling work?

People with self-control problems: driven by short-term utility (see Figure 1(A))

- ○ "I wanna enjoy music now (but wasted time is ignored)"

- ○ "I do not wanna exercise now (but potential benefit is underestimated)"

Figure 1 Theoretical Inverted Shapes of the Utility Streams Obtained from Engaging in (A) Wants and (B) Shoulds, Which Are Strategically Combined by Temptation Bundling
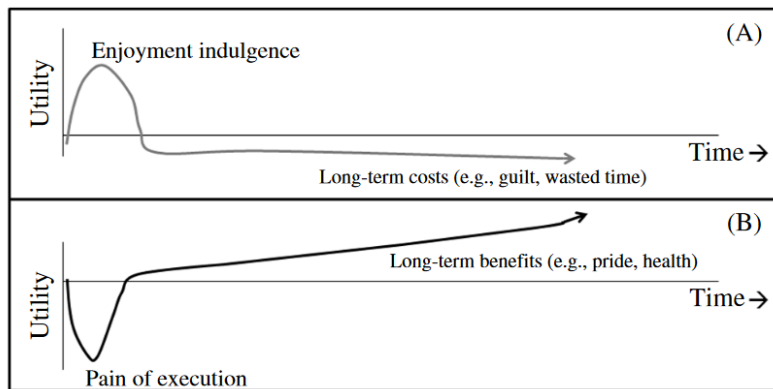
"Want+Should" > "Should"

Let's make a pact: I can **enjoy music(=want) only when exercising(=should)**

Choices (Before)

For short-sighted people, the choices before temptation bundling is:

{exercise, music, {exercise+music} }

The enjoying music only will be chosen since its highest short-term happiness.

Choices (After)

after restricting yourself from choosing {music}, the available choices now become:

{exercise, ~~{music}~~, {exercise+music} },

which makes {exercise+music} more desirable.
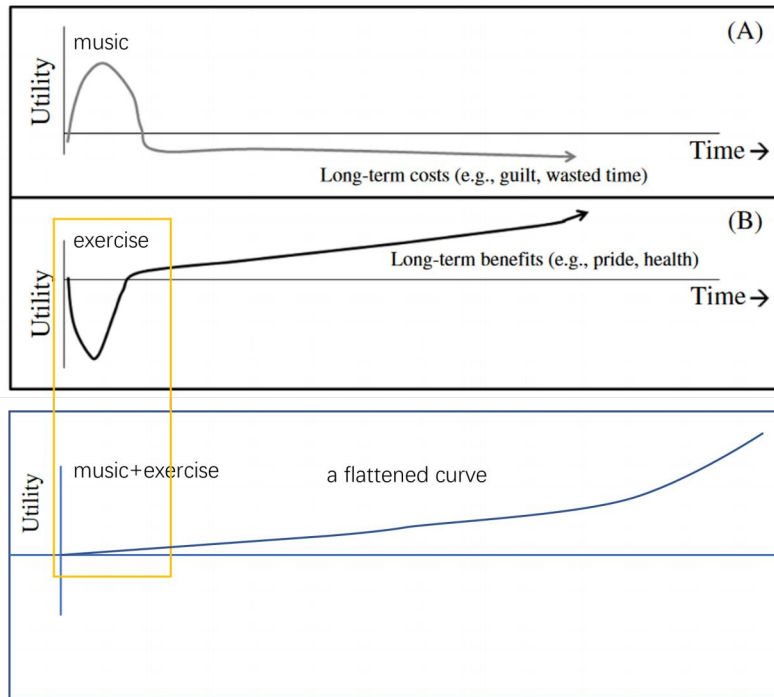
Theoretical Development

Few "should" behaviors? Increased regret and wasted time?
solve the negative effects together!

- valuable, healthy behaviors could be increased

- guilt and wasted time from indulgent behaviors can be decreased

Theoretical Development

New features in Temptation Bundling

- do not require monetary transfers (or any other form of punishment)
- temptation bundling may be particularly effective if complementarities exist between a temptation item and the healthy behavior it is bundled with
- the extent that the tempting activities bundled with should behaviors are habit forming, they might serve as powerful motivators

Experimental Design

Experiment

Goal

- **behavior change**: test whether temptation bundling can induce more effort
  - test whether individuals are able to effectively self-impose a suggested temptation bundling rule.
- **willingness to pay**: whether individuals intend to restrict their own behavior to garner the benefits of temptation bundling

Experiment

Treatments: three-conditioned and randomized

The experiment is conducted in an university's fitness gym (226 students, faculty, and staff; 10-week exercise study for 100$)

Condition	Items	Accessibility	Temptation Bundle
full treatment	an iPod containing 4 wanted audio novels	only at gym	strong
intermediate treatment	an iPod containing 4 want audio novels	at any time	mediate
control treatment	$25 (= value of novel books)	no	weak

The intermediate intervention is to exclude a potential issue that the effectiveness of this method comes from creating a structured environment that limits participants' ability to behave myopically.

Procedures

1. prescreening online survey to confirm eligibility (478)

2. randomly assigned with **one of the treatments**. (226/478)

3. online survey (fundamental data; choose wanted audio novels except people in control)

4. **Intake Visit** (215/226; collected biometric data and online survey data in behavioral lab)

   - In online survey: people in control is reminded of health, and for people in full/intermediate:

In order to make each workout you engage in at [university gym's name] gym more tempting (so that after a long day, you will actually find yourself craving a workout rather than dreading one), we recommend that you try making a rule for yourself: *only let yourself enjoy these novels while exercising.*

The hope is that you will frequently find yourself longing to find out what happens next in your novel, and this will lead to trips to the gym to satisfy that curiosity. Before long, we hope you will find yourself exercising more regularly as a result of your audio-novel addiction.

Full treatment

In order to make each workout you engage in at [university gym's name] gym more tempting (so that after a long day, you will actually find yourself craving a workout rather than dreading one), *we will only let you enjoy these novels while exercising at [university gym's name].* The iPod we are loaning you during this study will be required to remain in a locker at [university gym's name] at all times when you are not checked into the gym facility, and study administrators will check the locker regularly to ensure compliance with the study protocols.
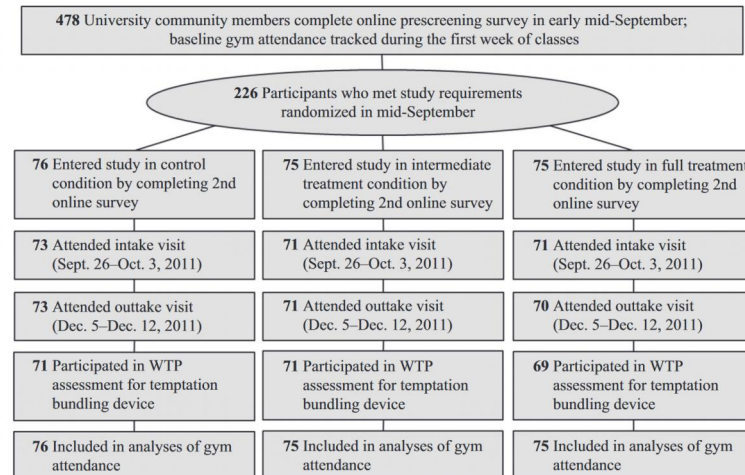
The hope is that you will frequently find yourself longing to find out what happens next in your novel, and this will lead to trips to the gym to satisfy that curiosity. Before long, we hope you will find yourself exercising more regularly as a result of your audio-novel addiction.

Intermediate treatment

Procedures

5. Nine-week exercise

6. **Outtake Visit** (214/215; biometric data again, online survey about exercise experience in lab)

7. **Willingness to Pay** assessment of Temptation bundle method

   ○ see details in the following section

8. Included in analysis of gym attendance

Figure 2   Flow of Study Participants

478 University community members complete online prescreening survey in early mid-September; baseline gym attendance tracked during the first week of classes

226 Participants who met study requirements randomized in mid-September

76 Entered study in control condition by completing 2nd online survey	**75** Entered study in intermediate treatment condition by completing 2nd online survey	**75** Entered study in full treatment condition by completing 2nd online survey
73 Attended intake visit (Sept. 26–Oct. 3, 2011)	**71** Attended intake visit (Sept. 26–Oct. 3, 2011)	**71** Attended intake visit (Sept. 26–Oct. 3, 2011)
73 Attended outtake visit (Dec. 5–Dec. 12, 2011)	**71** Attended outtake visit (Dec. 5–Dec. 12, 2011)	**70** Attended outtake visit (Dec. 5–Dec. 12, 2011)
71 Participated in WTP assessment for temptation bundling device	**71** Participated in WTP assessment for temptation bundling device	**69** Participated in WTP assessment for temptation bundling device
76 Included in analyses of gym attendance	**75** Included in analyses of gym attendance	**75** Included in analyses of gym attendance

Note. WTP, willingness to pay.

Experimental Result

Effectiveness of Intervention

- **self-reported biometric data**: control

- **"should" behavior**: attendance at gym

- **"want" behavior**: listening to 4 wanted audiobooks

- **bundle** = the access to audiobooks while exercising

Individual characteristics

No significant differences in individual characteristics

Table 1 **Pretreatment Characteristics of the Study Sample Reveal No Significant Differences Between Groups in Pretreatment Exercise Frequency, BMI, or Gender**

	All ($N = 226$)	Control group ($N = 76$)	Intermediate treatment group ($N = 75$)	Full treatment group ($N = 75$)
Visits to gym in first week of school	1.5 (1.6)	1.5 (1.8)	1.4 (1.5)	1.4 (1.6)
Self-reported weekly minutes of exercise	100.2 (66.6)	106.1 (69.1)	98.0 (65.5)	96.3 (65.6)
BMI (based on self-reported height and weight)	23.2 (3.9)	23.2 (4.3)	23.3 (4.0)	23.0 (3.4)
Male (%)	34.3	36.8	33.3	32.0

Note. Standard deviations in parentheses.

Average Weekly Gym Attendance (pre-Thanksgiving)

○ average total gym visits: 7.8 in full, 6.5 in intermediate, and 6.1 in control

○ average percentage of participants who visited the gym at least once in a given week: 51% in full, 44% in intermediate, and 42% in control

○ non-parametric tests: significance lies between full and control, intermediate and control

　○ (H0: full treatment = control, p = 00075; H0: intermediate treatment = control, p > 0010)

Weekly attendance among treatments
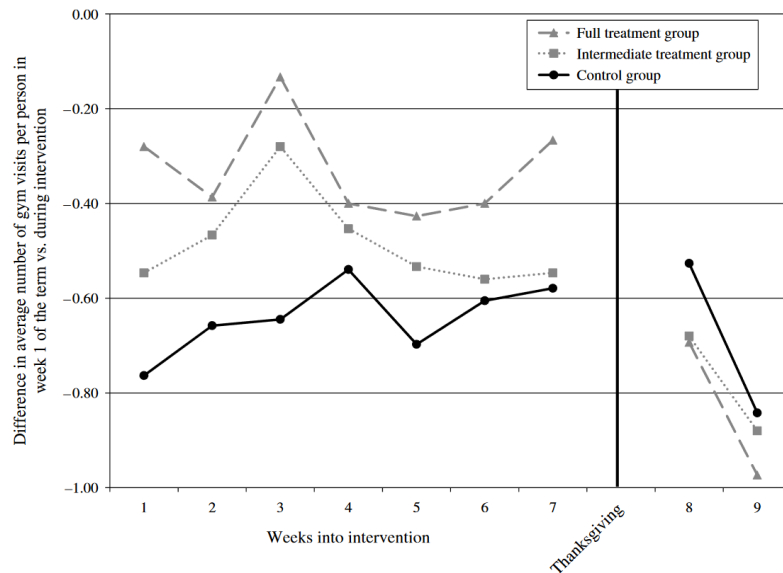
Average Weekly Gym Attendance (pre-Thanksgiving)

Regression indicates a difference between

- full treatment condition and the control condition of 0.31 gym visits per week (p = 0.026)
- intermediate treatment condition and the control condition of 0.14 gym visits per week (p > 0.10)

Regression results (partial)

Table 3   Ordinary Least Squares Regressions Predicting Weekly Gym Attendance, with Robust Standard Errors Clustered by Participant

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Primary predictor variables							
Full treatment	0.31*	0.21	0.48**	0.47**	0.46**	0.53**	0.50**
	(0.14)	(0.13)	(0.17)	(0.17)	(0.16)	(0.17)	(0.17)
Intermediate treatment	0.14	0.09	0.27†	0.27	0.28†	0.27†	0.28†
	(0.14)	(0.13)	(0.16)	(0.17)	(0.16)	(0.17)	(0.17)
Weeks since intervention × Full treatment	—	—	−0.07**	−0.06**	−0.07**	−0.07**	−0.07**
	—	—	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Weeks since intervention × Intermediate treatment	—	—	−0.05†	−0.04†	−0.05†	−0.04†	−0.04†
	—	—	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)

(Table 3, Model 1, cluster at participant-level)

Average Weekly Gym Attendance (post-Thanksgiving)

- A "reversed" pattern compared to habits in pre-Thanksgiving period.
- consistent with past research that exercise habits wear off over time and especially precipitously during holiday breaks

(Acland and Levy 2013)
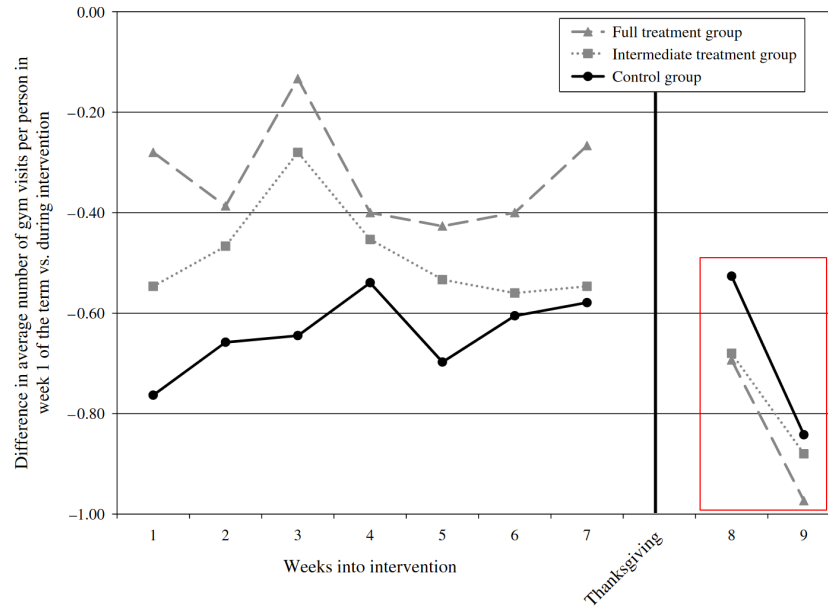
Average Weekly Gym Attendance (post-Thanksgiving)

- **Variables**: add a control for the number of weeks since intake visit and an interaction between this variable and each of our treatment indicators

- **Regression-adjusted difference increases**:
  - full-control condition is 0.48 visits per week(p = 0.004, 51% increase)
  - intermediate-control condition is 0.27 visits per week (p = 0.092, 29% increase)

- **Treatment effect over time**:
  - full treatment effect: decrease by 0.07 over weeks (p = 0.005)
  - control treatment effect: decrease by 0.05 over weeks (p = 0.069)

Table 3   Ordinary Least Squares Regressions Predicting Weekly Gym Attendance, with Robust Standard Errors Clustered by Participant

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Primary predictor variables							
Full treatment	0.31*	0.21	0.48**	0.47**	0.46**	0.53**	0.50**
	(0.14)	(0.13)	(0.17)	(0.17)	(0.16)	(0.17)	(0.17)
Intermediate treatment	0.14	0.09	0.27†	0.27	0.28†	0.27†	0.28†
	(0.14)	(0.13)	(0.16)	(0.17)	(0.16)	(0.17)	(0.17)
Weeks since intervention × Full treatment	—	—	−0.07**	−0.06**	−0.07**	−0.07**	−0.07**
	—	—	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Weeks since intervention × Intermediate treatment	—	—	−0.05†	−0.04†	−0.05†	−0.04†	−0.04†
	—	—	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)

(Table 3, Model 3)

Experimental Result

Treatment Effect Interactions

- **intake workout enjoyment**: self-reported enjoyment of the workout

- **availability**: proxy for a how busy a participant is

Workout enjoyment and Attendance

- a one-standard-deviation increase in self-reported intake visit workout enjoyment corresponds to an increase in the initial treatment effect of 0.21 gym visits per week.

Table 3 Ordinary Least Squares Regressions Predicting Weekly Gym Attendance, with Robust Standard Errors Clustered by Participant

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Primary predictor variables							
Full treatment	0.31*	0.21	0.48**	0.47**	0.46**	0.53**	0.50**
	(0.14)	(0.13)	(0.17)	(0.17)	(0.16)	(0.17)	(0.17)
Intermediate treatment	0.14	0.09	0.27†	0.27	0.28†	0.27†	0.28†
	(0.14)	(0.13)	(0.16)	(0.17)	(0.16)	(0.17)	(0.17)
Weeks since intervention × Full treatment	—	—	−0.07**	−0.06**	−0.07**	−0.07**	−0.07**
	—	—	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Weeks since intervention × Intermediate treatment	—	—	−0.05†	−0.04†	−0.05†	−0.04†	−0.04†
	—	—	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)
Moderators							
Z-availability × Full treatment	—	—	—	—	−0.21*	—	−0.22**
	—	—	—	—	(0.08)	—	(0.08)
Z-availability × Intermediate treatment	—	—	—	—	−0.07	—	−0.11
	—	—	—	—	(0.09)	—	(0.09)
Z-availability	—	—	—	—	−0.02	—	0.01
	—	—	—	—	(0.05)	—	(0.05)
Z-intake workout enjoyment × Full treatment	—	—	—	—	—	0.21*	0.21*
	—	—	—	—	—	(0.10)	(0.10)
Z-intake workout enjoyment × Intermediate treatment	—	—	—	—	—	0.06	0.08
	—	—	—	—	—	(0.13)	(0.13)
Z-intake workout enjoyment	—	—	—	—	—	0.07	0.07
	—	—	—	—	—	(0.06)	(0.06)

(p = 00041; Table 3, Models 6 and 7).

Workout enjoyment and Attendance

○ a one-standard-deviation increase in self-reported intake visit workout enjoyment corresponds to an increase in the initial treatment effect of 0.21 gym visits per week.

Table 3   Ordinary Least Squares Regressions Predicting Weekly Gym Attendance, with Robust Standard Errors Clustered by Participant

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Primary predictor variables							
Full treatment	0.31*	0.21	0.48**	0.47**	0.46**	0.53**	0.50**
	(0.14)	(0.13)	(0.17)	(0.17)	(0.16)	(0.17)	(0.17)
Intermediate treatment	0.14	0.09	0.27†	0.27	0.28†	0.27†	0.28†
	(0.14)	(0.13)	(0.16)	(0.17)	(0.16)	(0.17)	(0.17)
Weeks since intervention × Full treatment	—	—	−0.07**	−0.06**	−0.07**	−0.07**	−0.07**
	—	—	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Weeks since intervention × Intermediate treatment	—	—	−0.05†	−0.04†	−0.05†	−0.04†	−0.04†
	—	—	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)
Moderators							
Z-availability × Full treatment	—	—	—	—	−0.21*	—	−0.22**
	—	—	—	—	(0.08)	—	(0.08)
Z-availability × Intermediate treatment	—	—	—	—	−0.07	—	−0.11
	—	—	—	—	(0.09)	—	(0.09)
Z-availability	—	—	—	—	−0.02	—	0.01
	—	—	—	—	(0.05)	—	(0.05)
Z-intake workout enjoyment × Full treatment	—	—	—	—	—	0.21*	0.21*
	—	—	—	—	—	(0.10)	(0.10)
Z-intake workout enjoyment × Intermediate treatment	—	—	—	—	—	0.06	0.08
	—	—	—	—	—	(0.13)	(0.13)
Z-intake workout enjoyment	—	—	—	—	—	0.07	0.07
	—	—	—	—	—	(0.06)	(0.06)

(p = 00041; Table 3, Models 6 and 7).

Workout enjoyment and Attendance

○ These findings are in line with our theory whereby the act of exercising is itself less unpleasant when experienced in combination with an audio novel

○ They further support the idea that the more pleasant participants find exercising with audio novels, the more benefits they experience from temptation bundling

Table 3 Ordinary Least Squares Regressions Predicting Weekly Gym Attendance, with Robust Standard Errors Clustered by Participant

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Primary predictor variables							
Full treatment	0.31*	0.21	0.48**	0.47**	0.46**	0.53**	0.50**
	(0.14)	(0.13)	(0.17)	(0.17)	(0.16)	(0.17)	(0.17)
Intermediate treatment	0.14	0.09	0.27†	0.27	0.28†	0.27†	0.28†
	(0.14)	(0.13)	(0.16)	(0.17)	(0.16)	(0.17)	(0.17)
Weeks since intervention × Full treatment	—	—	−0.07**	−0.06**	−0.07**	−0.07**	−0.07**
	—	—	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Weeks since intervention × Intermediate treatment	—	—	−0.05†	−0.04†	−0.05†	−0.04†	−0.04†
	—	—	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)
Moderators							
Z-availability × Full treatment	—	—	—	—	−0.21*	—	−0.22**
	—	—	—	—	(0.08)	—	(0.08)
Z-availability × Intermediate treatment	—	—	—	—	−0.07	—	−0.11
	—	—	—	—	(0.09)	—	(0.09)
Z-availability	—	—	—	—	−0.02	—	0.01
	—	—	—	—	(0.05)	—	(0.05)
Z-intake workout enjoyment × Full treatment	—	—	—	—	—	0.21*	0.21*
	—	—	—	—	—	(0.10)	(0.10)
Z-intake workout enjoyment × Intermediate treatment	—	—	—	—	—	0.06	0.08
	—	—	—	—	—	(0.13)	(0.13)
Z-intake workout enjoyment	—	—	—	—	—	0.07	0.07
	—	—	—	—	—	(0.06)	(0.06)

Availability and Attendance

○ temptation bundling particularly benefits our busiest participants

   ○ availability = how busy a person is

   ○ a onestandard-deviation decrease in availability corresponds to an increase in the initial treatment effect of 0.21 gym visits per week (p = 0.014)

Table 3 Ordinary Least Squares Regressions Predicting Weekly Gym Attendance, with Robust Standard Errors Clustered by Participant

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Primary predictor variables							
Full treatment	0.31*	0.21	0.48**	0.47**	0.46**	0.53**	0.50**
	(0.14)	(0.13)	(0.17)	(0.17)	(0.16)	(0.17)	(0.17)
Intermediate treatment	0.14	0.09	0.27†	0.27	0.28†	0.27†	0.28†
	(0.14)	(0.13)	(0.16)	(0.17)	(0.16)	(0.17)	(0.17)
Weeks since intervention × Full treatment	—	—	−0.07**	−0.06**	−0.07**	−0.07**	−0.07**
	—	—	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Weeks since intervention × Intermediate treatment	—	—	−0.05†	−0.04†	−0.05†	−0.04†	−0.04†
	—	—	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)
Moderators							
Z-availability × Full treatment	—	—	—	—	−0.21*	—	−0.22**
	—	—	—	—	(0.08)	—	(0.08)
Z-availability × Intermediate treatment	—	—	—	—	−0.07	—	−0.11
	—	—	—	—	(0.09)	—	(0.09)
Z-availability	—	—	—	—	−0.02	—	0.01
	—	—	—	—	(0.05)	—	(0.05)
Z-intake workout enjoyment × Full treatment	—	—	—	—	—	0.21*	0.21*
	—	—	—	—	—	(0.10)	(0.10)
Z-intake workout enjoyment × Intermediate treatment	—	—	—	—	—	0.06	0.08
	—	—	—	—	—	(0.13)	(0.13)
Z-intake workout enjoyment	—	—	—	—	—	0.07	0.07
	—	—	—	—	—	(0.06)	(0.06)

(Table 3, Models 5 and 7).

Experimental Result

Willingness to Pay for Temptation Bundling Devices

Willingness to pay: details and result

A chance of winning an iPod

○ All 212 participants who completed a study outtake visit were told they had a chance of winning an iPod Shuffle loaded with one audio novel of their choice,
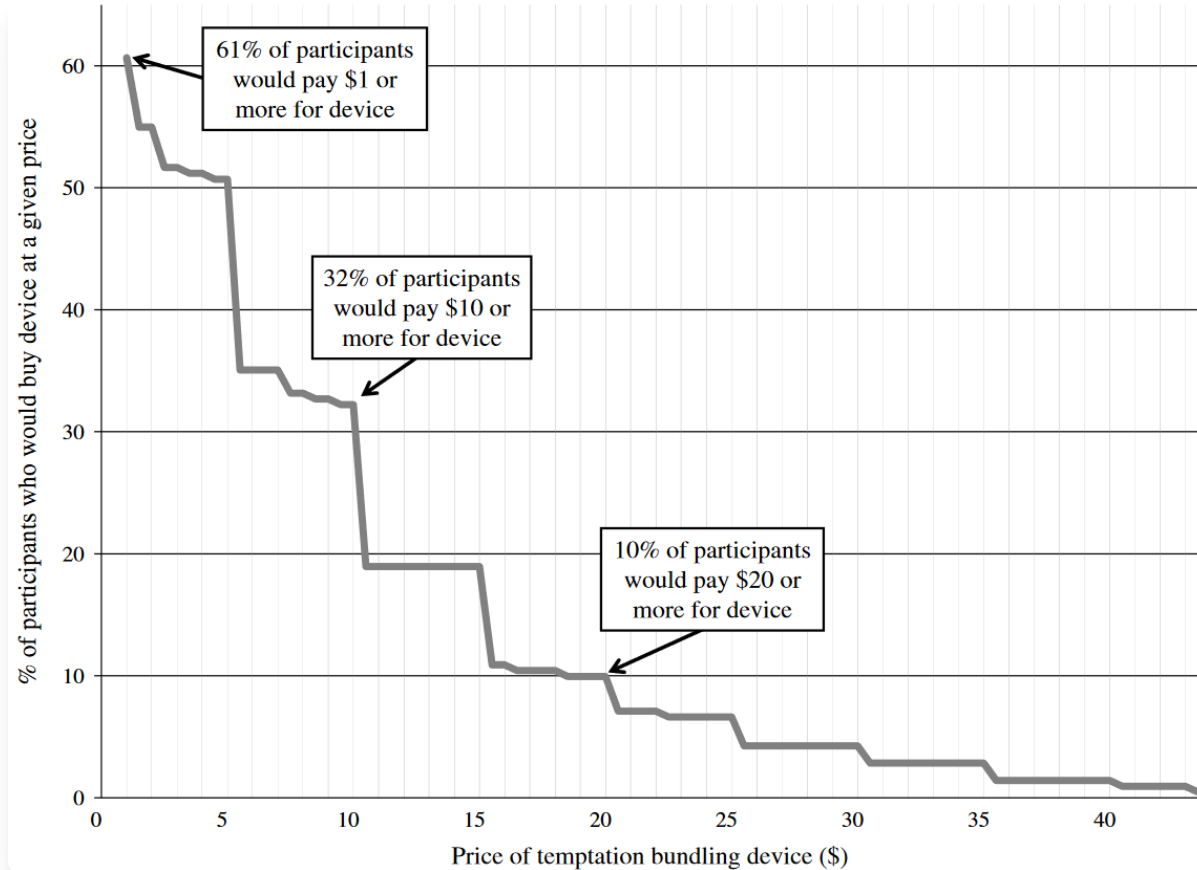
An opportunity to pay for an exercise program

○ if participants say yes, study staff would hold their iPod for one month in a monitored locker at the university gym, ensuring they never listened at home and had something to look forward to during exercise

○ 64% of participants indicated "sound[ed] appealing" on a "yes/no" scale ($t(211) = 19.43$; $p < 0.0001$)

Degroot-Marcshak reservation price elicitation

○ test the willingness to pay

○ Average willingness to pay for the program was $6.91 ($t(210) = 9.34$; $p < 0.0001$), with only 39.3% ($t(210) = 18.00$; $p < 0.0001$) of participants stating a $0 willingness to pay

(Becker et al. 1964)

Percentage of Study Participants Willing to Pay for a Temptation Bundling Device at a Given Price

61% of participants would pay $1 or more for device

32% of participants would pay $10 or more for device

10% of participants would pay $20 or more for device

% of participants who would buy device at a given price

Price of temptation bundling device ($)

Evidence: sophisticated decisions about pay for the program

○ the effectiveness of the intermediate treatment was significantly lower for those who expressed a **non-zero willingness to pay** for the program (β wtp>0_x_intermediate = −0.59; p = 0.013)

○ In other words, the better self-imposed temptation bundling rules worked for a participant, the less likely she was to place value on the externally monitored temptation bundling program

Discussion Questions

Next Class